Dispersal Dynamics in Food Webs

Melián, Carlos J ; Křivan, Vlastimil ; Altermatt, Florian ; Starý, Petr ; Pellissier, Loïc ; De Laender, Frederik

Abstract: Studies of food webs suggest that limited nonrandom dispersal can play an important role in structuring food webs. It is not clear, however, whether density-dependent dispersal fits empirical patterns of food webs better than density-independent dispersal. Here, we study a spatially distributed food web, using a series of population-dispersal models that contrast density-independent and density-dependent dispersal in landscapes where sampled sites are either homogeneously or heterogeneously distributed. These models are fitted to empirical data, allowing us to infer mechanisms that are consistent with the data. Our results show that models with density-dependent dispersal fit the , , and tritrophic richness observed in empirical data best. Our results also show that density-dependent dispersal leads to a critical distance threshold beyond which site similarity (i.e., tritrophic richness) starts to decrease much faster. Such a threshold can also be detected in the empirical data. In contrast, models with density-independent dispersal do not predict such a threshold. Moreover, preferential dispersal from more centrally located sites to peripheral sites does not provide a better fit to empirical data when compared with symmetric dispersal between sites. Our results suggest that nonrandom dispersal in heterogeneous landscapes is an important driver that shapes local and regional richness (i.e., and tritrophic richness, respectively) as well as the distance-decay relationship (i.e., tritrophic richness) in food webs.

DOI: https://doi.org/10.1086/679505

Multi-trophic metacommunities and the biogeography of food webs

Carlos J. Melián[1,2,*,**], Vlastimil Křivan[3,4,**], Florian Altermatt[5],
Peter Starý[3], Loïc Pellissier[6], and Frederik De Laender[7]

[1]Fish Ecology and Evolution Department,
Center for Ecology, Evolution and Biogeochemistry,
Swiss Federal Institute of Aquatic Science and Technology, Switzerland.
[2]National Center for Ecological Analysis and Synthesis
University of California at Santa Barbara, USA
[3]Biology Centre ASCR, Institute of Entomology, České Budějovice, Czech Republic.
e-mail:vlastimil.krivan@gmail.com and stary@entu.cas.cz
[4]Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.
[5]Department of Aquatic Ecology,
Swiss Federal Institute of Aquatic Science and Technology, Switzerland.
e-mail:florian.altermatt@eawag.ch
[6]The Arctic Research Centre, Department of Bioscience,
Aarhus University, Aarhus, Denmark. e-mail:loic.pellissier@unil.ch
[7]Research Unit in Environmental and Evolutionary Biology,
Université de Namur ASBL, Belgique,
e-mail: frederik.delaender@unamur.be

Keywords: α, β and γ-tritrophic richness,
distance-decay relationship, metacommunity dynamics,
spatial food webs, heterogeneous landscape.

Type of Article: Article

Number of figures: 3 (1 and 2 in color); Number of tables: 4
* Corresponding author:
Swiss Federal Institute of Aquatic Science and Technology, Seestrasse 79,
Kastanienbaum CH-6047, Switzerland.
Phone: +41 58 765 2208; Fax: +41 58 765 2168; e-mail: carlos.melian@eawag.ch.
** Shared first authorship

Abstract

Patterns of trophic richness in an empirical multi-trophic metacommunity consisting of plants, aphids and their parasitoids are analyzed. Empirical observations are compared with predictions of a series of models that differ in dispersal (density-independent vs. density-dependent; symmetric vs. asymmetric) and spatial site distributions (homogeneous vs. heterogeneous landscapes). Among these models, the model with density-dependent and symmetric dispersal in heterogeneous landscapes fits $\alpha-$, $\beta-$ and $\gamma-$tritrophic richness observed in empirical data best. This model predicts a critical distance threshold between sites below which two sites are quite similar but beyond which the decrease in similarity is much faster. Such a threshold is also observed in the empirical data. Our results suggest that density-dependent dispersal in heterogeneous landscapes is one of the key drivers connecting $\alpha-$tritrophic richness to turnover rates across geographically distant food webs.

1 Introduction

Species coexistence in metacommunities (i.e., groups of spatially structured communities connected by dispersal, Leibold et al., 2004), depends both on trophic interactions at each local community and on patterns of dispersal between these communities (Araújo and Luoto, 2007; Boulangeat et al., 2012). Metapopulation theory (Levins, 1962; Hanski, 1999) predicts that despite local extinctions, a population can survive in a fragmented landscapes consisting of several patches. In metacommunities with two trophic levels intermediate dispersal rates can stabilize otherwise unstable resource-consumer population dynamics (Huffaker, 1958). Theoretical research (e.g. Murdoch et al., 2003; Briggs and Hoopes, 2004; Křivan, 2008) showed that necessary conditions for such global species coexistence in resource-consumer metacommunities are either differences in local population dynamics, or differences in dispersal dynamics between patches. Moreover, dispersal rates cannot be neither too low (because low dispersal rates do not rescue populations from global extinction), nor too high to avoid synchronization of population dynamics (Gouhier et al., 2010). However, for metacommunities with resource-consumer-predator food chain in each patch Koelle and Vandermeer (2005) showed that an increase in the dispersal rates can reduce synchrony in population dynamics. In multi-trophic communities a peak in food web complexity and species diversity arises for intermediate dispersal rates (Pillai et al., 2011).

Most of the theoretical work on metacommunities assumes density independent (e.g., random) dispersal between patches. Theoretical work on simple di- and tri-trophic metacommunities shows that density-dependent dispersal in direction of a higher fitness promotes species coexistence by weakening competition (Holt and Hoopes, 2005; Křivan, 2014). Yet, it is not clear which dispersal models (i.e., density-independent or density-dependent dispersal dynamics) predict better the empirical patterns of spatial food webs

(Koelle and Vandermeer, 2005; Amarasekare, 2008; Rezende et al., 2009; Massol et al., 2011; Thuiller et al., 2013). This knowledge gap is particularly relevant when confronting dispersal models with extensive datasets on geographically distant food webs (Smith et al., 2011; Massol et al., 2011; Kissling et al., 2012).

Metacommunities are often characterized by three richness indexes: $\alpha-$richness measures local richness (i.e., the number of distinct species at a site), $\gamma-$richness measures regional richness (i.e., the number of distinct species at the region/landscape), and $\beta-$richness measures changes in community composition by comparing species composition at two or more sites. A measure that is particularly relevant for describing metacommunities is the distance-decay relationship in community similarity (Nekola and Peter, 1999; Morlon et al., 2008) expressed as the dependence of the $\beta-$richness on geographic distance between sites. It is known that the distance-decay is influenced by several factors, e.g., spatial organization of communities, local species abundances and population aggregation (Morlon et al., 2008). Yet, the underlying mechanisms generating the distance-decay in metacommunities remain poorly understood (Bolker, 2004; Dunne, 2006; Massol et al., 2011; Poisot et al., 2012).

In this article we focus on the effect of dispersal on metacommunity dynamics. We analyze empirical data describing a multi-trophic metacommunity consisting of plants, aphids, and their parasitoids (Starý, 2006). First, we develop several models of tritrophic metacommunities. These models focus on random vs. nonrandom dispersal, symmetric vs. asymmetric dispersal, and homogeneous vs. heterogeneous spatial distribution of sites. Second, using the empirical data we parametrize these models which allows us to select the model that fits the data best. We show theoretically that density-dependent dispersal in heterogeneous landscapes has a strong influence on β-tritrophic richness measured as the similarity between two or three sites. In particular, we show that under density-dependent dispersal a critical distance exists, beyond which similarity between

two sites sharply decreases. Such critical distance is also observed in the empirical data. Models with random dispersal do not predict such a threshold. Moreover, models with density-dependent dispersal predict steeper decrease in β-tritrophic richness with distance when compared to models with random dispersal. Again, we show that such trends better agree with empirical observations. Our results suggest that nonrandom dispersal in heterogeneous landscapes is an important driver that shapes local (i.e., α-tritrophic richness) and regional richness (i.e., γ-tritrophic richness) as well as the distance-decay relationship in multi-trophic metacommunities (i.e., β-tritrophic richness).

2 Plant-aphid-parasitoid data

The data that we analyze in this article describe tritrophic associations between 411 plant species, 267 aphid species, and 302 Hymenoptera parasitoid species (family Braconidae and subfamily Aphidiinae, Starý, 2006). The data were collected at 302 sites in the Czech Republic between 1954 and 2004. Each site is characterized by two coordinates (x_1, x_2) corresponding to its position on a grid overlayed on the map of the Czech republic. The coordinates allow us to calculate the distance between two sites (x_1, x_2) and (y_1, y_2) as $d_{ij} = \sqrt{A^2(x_1 - y_1)^2 + B^2(x_2 - y_2)^2}$, where A = 12km and B = 11.1km is the grid size of 12 $\times$ 11.1km^2. To capture the heterogeneity in site distribution across the landscape we calculate for each site its total geographic distance from other sites (i.e., the sum of all distances between the focal site and all other sites). The corresponding distribution is used to classify landscapes as homogeneous or heterogeneous. If distances between sites are independent randomly distributed variables, the sum converges to a normal distribution. In what follows we call landscapes with normally distributed geographic distance homogeneous landscapes while those where distribution significantly deviates from the normal distribution are called heterogeneous landscapes.

Sampling effort varied among sites as some were sampled many times, while others only

once. In this article we count each observed chain at a given site only once, i.e. multiple reports of one chain at a given location are not accounted for in this analysis. For 229 sites, the corresponding habitat type was reported with some sites containing multiple habitat types (i.e., undergrowth, field, ruderal, road, hedge, park, steppe, meadow, pond, garden, forest, deciduous trees, gravel, orchard, town, greenhouse, alley, waste, rocky, and grassland; Starý, 2006).

3 $\alpha-$, $\beta-$, and $\gamma-$richness of tritrophic chains

To analyze tritrophic associations between plants, herbivores, and parasitoids, we adjust standard richness indexes. We define $\alpha-$tritrophic richness as the total number of different tritrophic chains in each sampled site. Two chains are different if they differ at least in one species. This measure is conceptually identical to $\alpha-$richness in ecological communities, but it considers tritrophic associations instead of species. Thus, $\alpha-$tritrophic richness is a measure of the local richness in food chain configurations. Similarly, the regional $\gamma-$tritrophic richness is measured as the number of unique tritrophic chains observed across all sites.

Classical $\beta-$richness in community ecology measures similarity between sites using pairwise comparison. This measure is useful when studying changes in species composition along ecological gradients. For two sites, $\beta-$richness compares species richness in each of the two sites and the number of species shared by the two sites (Magurran, 2004; Poisot et al., 2012). If the data is a random collection of samples from a large region, then, in addition to the pairwise comparison, a multiple-site similarity measure is required to better capture the heterogeneity of habitats (Diserud and Ødegaard, 2007). In our context we redefine this index based on tritrophic chains instead of species. We calculate the 2- and 3-site Sørensen similarity index (Diserud and Ødegaard, 2007) as $C_S^2 = 2c_{ij}/(u_i + u_j)$ and $C_S^3 = \frac{3}{2}(c_{ij} + d_{ik} + e_{jk} - f_{ijk})/(u_i + u_j + u_k)$, where u_i, u_j, and u_k are the observed

number of unique tritrophic chains in sites i, j and k, respectively, and c_{ij}, d_{ik}, e_{jk}, and f_{ijk} are the number of chains shared by i and j, i and k, j and k, and i, j and k sites, respectively. Tritrophic chain similarity varies between 0 (completely dissimilar, no trophic chains in common) and 1 (completely similar, all trophic chains shared).

In this article we are interested in dependence of $\beta-$tritrophic similarity on distance between sites, i.e., on the distance-decay of similarity. Instead of calculating the mean $\beta-$tritrophic similarity index at a given distance, we focus on the maximum $\beta-$tritrophic similarity index which better captures the decline in similarity between two (or three) sites (see Results section). In the case of the 2-site index we study this dependence as a function of distance between two sites, while in the case of the 3-site index we plot this index as a function of the mean distance between three sites. We also compute the 2-site habitat $\beta-$tritrophic similarity index conditioned on the same habitat type. For example, let us consider the situation where there are habitats types a, b, c in site A while site B contains only habitat types a and c. To compute the 2-site $\beta-$tritrophic similarity index, we use all tritrophic chains found at both sites A and B independently of the habitat types. For the 2-site habitat $\beta-$tritrophic similarity index we compare only those tritrophic chains in habitat a only, or habitat c only.

We tested the robustness of tritrophic richness indexes to sampling effort by studying how they change when the number of sample sites increases (Polis, 1991; Bersier and Sugihara, 1999). We did this by randomly sampling an increasing number of sites taking into account all the unique tritrophic chains observed in each site, starting by 5, 10, 20, 50, 100, 200, and 300 sites from the original dataset and calculating all three richness measures for each of these subsets. Because these measures were quantified as distributions, we described each measure using its mean value and its standard deviation. Random samplings were repeated 1000 times to account for variability in tritrophic richness among sites.

4 Models

We consider a landscape consisting of several sites. At each site there is a food web consisting of resources (R), consumers (H), and parasitoids (P). The food chains at all sites represent a tritrophic metacommunity. To model spatio-temporal changes in population abundances we need to define population and dispersal dynamics (the key terms are summarized in Table 1). To simplify population dynamics we assume that the number of individuals at each trophic level at a given site is fixed and equals to the site environmental carrying capacity for the given trophic level. Thus, the overall number of individuals at each trophic level is fixed, but species composition changes in time due to replacement of a dead individual by an individual of possibly another species within the same trophic level.

4.1 Model 1: Density- and site-independent dispersal dynamics

The first model assumes that dispersal rates are inversely proportional to the geographic distance between sites. This leads to the dispersal rate of species k_ϕ in metacommunity ϕ from site j to site i (where ϕ stands either for the resource (R), the consumer (H), or the parasitoid metacommunity (P))

$$m_{ij}^{k_\phi} = \frac{m_\phi}{d_{ij}}. \tag{1}$$

Here d_{ij} is the geographical distance between site i and j, and m_ϕ is the intensity of emigration rate specific for each metacommunity ϕ. Because dispersal from site i to site j is the same as in the opposite direction ($m_{ji}^{k_\phi} = m_{ij}^{k_\phi}$), model 1 represents symmetric and site- and density-independent dispersal where dispersal to less distant sites is more likely than dispersal to more distant sites.

4.2 Model 2: Site-dependent dispersal dynamics

The second model assumes that dispersal rates depend not only on the geographic distance but also on the environmental carrying capacity (denoted as J_i^{ϕ} for site i and metacommunity ϕ) of the receiving site. The dispersal rate from site j to site i of a species k_{ϕ} is as in model 1 ($m_{ij}^{k_{\phi}} = \frac{m_{\phi}}{d_{ij}}$), but now dispersing individuals have a higher probability to settle in a site that has a larger carrying capacity (which may be proportional to the site area). This leads to the following probability to settle in site i

$$p_i^{\phi} = \frac{J_i^{\phi}}{\sum_{j=1}^{N} J_j^{\phi}}, \tag{2}$$

where N is the number of sites. Thus, the dispersal rate of species k_{ϕ} in metacommunity ϕ from site j to site i is

$$m_{ij}^{k_{\phi}} = \frac{p_i^{\phi} m_{\phi}}{d_{ij}}.$$

Dispersal model 2 represents symmetric and density-independent dispersal. However, dispersal is site-dependent, because immigration to larger sites (i.e., sites with a larger environmental carrying capacity) is more likely when compared with smaller sites.

4.3 Model 3: Density-dependent dispersal dynamics

The third model considers that dispersal rates depend not only on the geographic distance but also on the number of individuals of the same species in the receiving site i. The dispersal rate of species k_{ϕ} in metacommunity ϕ from site j to site i is

$$m_{ij}^{k_{\phi}} = \frac{J_i^{\phi} - N_i^{k_{\phi}}}{J_i^{\phi}} \frac{m_{\phi}}{d_{ij}}, \tag{3}$$

where J_i^{ϕ} has the same meaning as in Model 2 and $N_i^{k_{\phi}}$ is the number of individuals of species k_{ϕ} of metacommunity ϕ that are already in site i. (We recall that the abundance of all individuals at a given trophic level is equal to the environmental carrying capacity of the

site, i.e. $\sum_{k_\phi} N_i^{k_\phi} = J_i^\phi$.) This means that the immigration rate to a given site decreases as the number of conspecifics already in the site increases. Such a negative density-dependent dispersal strategy often leads to the ideal free distribution (Fretwell and Lucas, 1970; Křivan et al., 2008). This may be the case where species have different niches so that they do not compete for their resources but there is competition for resources among individuals within the same species. Thus, this model considers intra-specific competition, but it neglects inter-specific competition. We note that equation (3) is the same as equation (1) when no conspecifics are present in the receiving site ($N_i^{k_\phi} = 0$). Dispersal model 3 represents symmetric and density-dependent dispersal because immigration depends negatively on the abundance of conspecifics at the receiving site.

4.4 Model 4: Density-dependent and asymmetric dispersal dynamics

The fourth model assumes that dispersal rates depend not only on the geographic distance and the number of conspecifics that are already in the receiving site, but also on the spatial distribution of sites. One particular realization of such asymmetric dispersal from site j to site i of a species k_ϕ is

$$m_{ij}^{k_\phi} = \begin{cases} \dfrac{J_i^\phi - N_i^{k_\phi}}{J_i^\phi} \dfrac{m_\phi}{d_{ij}} & \text{if } \sum_{k=1}^N d_{ik} \geq \sum_{k=1}^N d_{jk}, \\ 0 & \text{if } \sum_{k=1}^N d_{ik} < \sum_{k=1}^N d_{jk} \end{cases} \tag{4}$$

where $N_i^{k_\phi}$ is the number of individuals of species k_ϕ of metacommunity ϕ that are already in site i, m_ϕ is the intensity of emigration rate specific for each metacommunity ϕ and N is the number of sites. Formula (4) assumes that dispersal is unidirectional from sites with lower total geographic distance (where total geographic distance is defined as the sum over all distances between the focal site and all other sites) to sites with a larger total geographic distance. If core sites are defined as those that have lower total geographic distance than peripheral sites, then formula (4) implies that individuals move from the

core to the periphery of the landscape (Channell and Lomolino, 2000).

4.5 Dispersal dynamics from the regional species pool

In addition to the dispersal dynamics between sites, new species can emerge with very low probability (ν_ϕ for metacommunity ϕ) from the regional species pool. We consider an extremely diverse regional species pool at each trophic level, containing an infinite number of species. Because of the infinite number of species in the regional pool, we assume that every immigration event introduces a new species. Immigration of a new species corresponds to speciation in the context of metacommunity models (Vanpeteghem and Haegeman, 2010).

4.6 Multi-trophic metacommunity dynamics

Here we describe population-dispersal dynamics. At each time step one site for each metacommunity ϕ is chosen with probability p_i^ϕ. For dispersal models 1, 3, and 4 this selection is random (i.e., $p_i^\phi = 1/N$ where N is the number of sites), while model 2 assumes probability given by eq. (2). In each site a death event occurs in each chosen metacommunity ϕ. This death event is compensated for by a recruitment of a new individual so that the total number of individuals does not change during simulations. However, as the new individual can be of another species, the species composition changes in time. The recruitment is either due to local reproduction (birth), or immigration from another site, or immigration from the regional species pool. The key parameters used throughout the article and the equations of the model are summarized in Table 2 and in Appendix 1 in the online edition of the American Naturalist, respectively.

5 Simulations and parameter estimation

We numerically simulate the four population-dispersal models in either homogeneous landscapes, or heterogeneous landscapes. For each combination we run 100000 replicates. We assume that initially, each trophic level contains only one species. For each simulation run (i.e., a replicate), the number of generations was chosen at random from a uniform distribution, $\mathcal{U}[100, 30000]$. The number of individuals per site (i.e., the overall site environmental carrying capacity J_i^{ϕ}) within each trophic level ϕ was set to 1000 in model 1, 3 and 4. For model 2 the number of individuals per site varies across sites within each replicate and these numbers were chosen from a normal distribution, $\mathcal{N}[1000, 100]$ where 1000 is the mean and 100 is the variance.

We set all mortality rates equal to 1 (i.e., natural mortality rate for plants (μ^{k_R}) and parasitoids (μ^{k_P}), and mortality rate of aphids due to parasitation ($\alpha^{k_H k_P}$)). Rates of immigration from the regional species pool, ν_{ϕ}, and the intensity of emigration rate specific for each metacommunity, m_{ϕ}, were chosen at random from a uniform distribution, $\mathcal{U}[10^{-4}, 10^{-2}]$, and $\mathcal{U}[10^{-3}, 7 \times 10^{-1}]$, respectively. Local birth rates for each metacommunity were set $\lambda_{\phi} = 1 - \nu_{\phi} - m_{\phi}$ so that a new individual replacing the dead individual appears with certainty.

In our simulations we check whether the resulting food web at the end of the simulation mimics the empirical connectance of the food webs. To this end we compare the connectance of the simulated and the empirical food web. Connectance in each site i for the resource-host and for the host-parasitoid is defined as $C_i^{RH} = \frac{L_i^{RH}}{S_i^R \times S_i^H}$, and $C_i^{HP} = \frac{L_i^{HP}}{S_i^H \times S_i^P}$, respectively. Here L_i^{RH}, L_i^{HP}, S_i^R, S_i^H, and S_i^P are the number of interactions between plant-aphids and aphids-parasitoids, the number of plant, aphids and parasitoid species, respectively. After each simulation run (i.e., a replicate) we check whether the simulated connectance in each site i between plants-aphids and aphids-parasitoids is in

the empirical range of the minimum and maximum connectance values, (0.02-0.24). A replicate was removed if any of the sites is outside the empirical range. About 10% of replicates were removed.

For each simulation run (i.e., for each specific parameter choice), we calculate the $\alpha-$ and $\gamma-$tritrophic richness and compare it with the empirical values. We use an $\epsilon-$tolerance rejection algorithm for model choice within an approximate Bayesian computation framework to do these comparisons (Beaumont, 2010). This means that for each simulation run we calculate the misfit (for definition see formula (A-2) in Appendix 2) between observed and predicted tritrophic chains. A simulation run is assumed to provide a good fit with empirical community if the misfit is above the $\epsilon-$tolerance threshold (see Appendix 2 and Figure A1). This means that such simulation runs predict $\alpha-$ and $\gamma-$tritrophic richness values that are sufficiently close to the empirical values. Model parameters that correspond to simulations that meet the $\epsilon-$tolerance threshold are called the best fit parameters.

6 Results

The distribution of sampled sites (Figure 1, black bars) is significantly different from a normal distribution with the empirical mean (145 km) and variance (50 km) ($p < 0.0001$, Kolmogorov-Smirnov test, Figure 1, red bars). In our terminology, this means that the sampled sites form a spatially heterogeneous landscape.

6.1 $\alpha-$ and $\gamma-$tritrophic richness

Empirical data contain 4966 tritrophic chains at all sites. Observed $\gamma-$tritrophic richness (i.e., the number of unique tritrophic chains) is 1304 tritrophic chains. Our analysis shows that sampling effort was not sufficient to estimate $\gamma-$tritrophic richness because it does not level off (Appendix 3 and Figure A2, panels A and B). At many sites only one

tritrophic chain was observed and fewer than 16% of sites hosted more than 10 different chains. The mean number of tritrophic chains per site, the mean $\alpha-$tritrophic richness, equals to 16.4, and the spatial variation in the number of tritrophic chains equals to 6078. The sampling effort to estimate $\alpha-$tritrophic richness was sufficient because $\alpha-$tritrophic richness saturates after sampling 302 sites (Figure A2, panels C and D).

Using empirical data we estimated parameters of our four models for the empirical (heterogeneous) site distribution that predict best the local and regional tritrophic richness (i.e., $\alpha-$, and $\gamma-$tritrophic richness). Density-dependent and symmetric (model 3) or asymmetric (model 4) dispersal predict local and regional richness of chains best. In Table 3 and 4 this corresponds to Bayes factors larger than 2 (see also Appendix 2 and Table A1). Model 4 with preferential dispersal from more centrally located sites to peripheral sites does not provide a better fit to empirical data when compared with model 3 with symmetric dispersal (Bayes factors values lower than 0.5, Table 3 and 4).

These results are graphically presented in Figure 2 ($\alpha-$tritrophic richness in the left column and $\gamma-$tritrophic richness in the right column). Each dot corresponds to one simulation with randomly chosen parameters from given intervals (see Section "Simulations and parameter estimation"). For each simulation we calculated the corresponding mean and variance of $\alpha-$tritrophic richness, and $\gamma-$tritrophic richness for different intensities of emigration rates. Then we calculated the misfit (Appendix 2) between the empirical data and the simulation. We plotted the distribution of these misfits (see Figure A1) and calculated the misfit that corresponds to quantile 1% (see the horizontal line in Figure 2). This quantile corresponds to $\epsilon = -1338.5$. The empirically observed mean (16.4) and variance (6078) of $\alpha-$richness and $\gamma-$tritrophic richness (1304) are plotted as vertical lines in Figure 2. Dots that are above the horizontal line correspond to parameters that fit the observed data best.

While the true mean number of tritrophic chains per site compares well with the

predicted values generated by models 3 and 4 for low to medium dispersal rates (i.e., 16.4 vs. 13.6±6; Figure 2, left panels), the true spatial variation of the number of tritrophic chains per site strongly deviates from the predicted variance by all models (Figure 2, middle panels, the observed variance is shown as the vertical line). Only when dispersal is high, models 3 and 4 predict the variance correctly, but in this case the misfit values strongly deviate from the observed value (Figure 2, middle panels C, D).

Predicted $\gamma-$tritrophic richness for density-dependent models 3 and 4 that fits the observed food web best does not show a significant difference from the observed $\gamma-$tritrophic richness (Table 4, right column of Figure 2, panels C, D). In other words, preferential dispersal from core sites to peripheral sites did not significantly improve results of non-preferential dispersal. Similarly to the $\alpha-$tritrophic richness, most tritrophic chains were observed only once across all the sampling sites, while only about 1% of the observed chains were observed in more than 10 locations. However, these predictions should be interpreted with caution, because, as we have already mentioned, the sampling effort was insufficient to estimate $\gamma-$tritrophic richness (Figure A2, panels A, B).

6.2 β-tritrophic richness

To study $\beta-$tritrophic richness we plot the maximum site similarity in the true tritrophic plant-aphid-parasitoid data as a function of geographic distance (Figure 3, dots). For a given distance between two or three sites, this figure shows the maximum number of shared tritrophic chains among two (panel A top) or three sites (panel B top, x-axes represents the mean distance between the three sites). While panels A and B in A3 do not consider habitat types within a given site, Panel C compares only tritrophic chains within the same habitat type. The solid line (together with confidence intervals shown by dotted lines) in Figure 3A-C corresponds to the maximum similarity values predicted by the model with density-dependent and symmetric dispersal (model 3). We observe that

most empirical values (shown as dots) fit within the confidence interval for the two site (panel A top) and the three site (panel B top) but the fit is not as good once only the same habitat types are compared (panel C top). In the latter case 23% of the empirically observed data significantly deviate from the predicted values.

Figure 3 shows that model 3 predicts lower decrease in similarity at short distances (see the slope of the solid line for distances approx. between 0 and 200 km), followed by a steeper decrease (approx. between 200 and 350 km). Beyond 350 km, the 2-site and 3-site tritrophic chain similarity values are near to zero (Figures 3A and B). However, when the same habitat types are compared, even sites whose distance is 600 km can be highly similar (Figure 3C). This means that tritrophic chains are specific for each habitat type. The question is how to detect this threshold distance in empirical data. To test whether there is a critical distance threshold beyond which the similarity between sites sharply decreases we calculate the variance of site similarity as a function of geographical distance. We note that small variance in similarity index at a given distance means that this index does not change too much for sites separated by this distance. At distances where this index sharply increases, the variance should be maximal. The variance in site similarity in empirical data peaks at approx. 350km for the 2- and the 3-site metrics (Figure 3A-C, bottom).

The same pattern is predicted by model 3 when applied to a heterogeneous landscape (Figure A4, panel D). The predicted distance-decay of variance of similarity shows no peak across all the geographic distances for the model 1 with density-independent dispersal neither in homogeneous (Figure A3, panel C, red dots) nor in heterogeneous (panel C, black dots) landscapes. Results for model 2 and model 4 are similar to model 1 and model 3, respectively, and these results are not shown here. These results suggest that within the mechanisms we tested the following two conditions lead to the empirical patterns of decay of maximum similarity between sites: 1) individuals disperse preferentially to

sites where conspecifics are rare and 2) asymmetric dispersal does not improve the fit between observed and simulated communities when compared to symmetric dispersal between sites (i.e, when the probability of dispersal from site i to j is the same as from j to i independently whether the site is located more centrally, or on the periphery of the sampled area). The effect of empirical sampling effort on regional similarity, the $\beta-$tritrophic richness, levels off after approximately 100 sites were sampled (Figure A2, panels E, F). This suggests that the sampling effort when collecting empirical data for plant-aphid-parasitoid interactions was adequate to estimate $\beta-$tritrophic richness.

6.3 Estimated parameter values

We estimate from model 3 the intensity of emigration rates and the immigration rates from the regional species pool that best predict the empirical tritrophic richness at local and regional scales. The median emigration and immigration rates from the regional species pool were obtained from all simulation runs with misfits above quantile 3% of the distribution of misfits (Figure A1). Neither emigration rates ($p > 0.1$, Kolmogorov-Smirnov test for all pairwise comparisons with median value $m_R \sim m_H \sim m_P \sim 0.05$) nor immigration rates from the regional pool show significant differences across trophic levels ($p > 0.1$, Kolmogorov-Smirnov test, with median value $\nu_R \sim \nu_H \sim \nu_P \sim 0.0045$). These estimated values imply that around 5% of individuals disperse per site and per generation within the sampled area. This represents a moderately strong dispersal limitation within the parameter range used in simulations (orange dots in Figure 2 represent low intensity of emigration rates, $m_\phi \in [10^{-3}, 10^{-1}]$).

7 Discussion

In this article we study mechanisms generating $\alpha-$, $\beta-$ and $\gamma-$ tritrophic richness observed in an empirical metacommunity consisting of plants, aphids and their parasitoids.

We compare these empirical observations with predictions of a series of models that differ in dispersal (density-independent vs. density-dependent; symmetric vs. asymmetric) and spatial distribution of sampled sites (homogeneous vs. heterogeneous landscapes). Among these models, the models with density-dependent dispersal fit $\alpha-$, $\beta-$ and $\gamma-$tritrophic richness observed in empirical data better than models with density independent dispersal (cf. Models 3 and 4 with density dependent dispersal vs. Models 1 and 2 with density independent dispersal in Figure 2). Our study confirms that nonrandom dispersal has a strong effect on empirical patterns of food webs connected by dispersal (Shurin, 2001; Economo and Keitt, 2008; Holt and Hoopes, 2005; Massol et al., 2011; Carrara et al., 2012). Our analysis shows that the results for symmetric dispersal are as good as for asymmetric dispersal which assumes preferential dispersal from the more centrally located sites to sites on the periphery of the sampled area (cf. results for Model 3 with symmetric dispersal vs. Model 4 with asymmetric dispersal in Table 3 and 4). Our results also show that models with density-dependent dispersal in heterogeneous landscapes that mimic the empirical site distribution agree with the observed pattern of the distance-decay of similarity. Using the 2- and 3-site Sørensen similarity index, we calculate the maximum number of shared tritrophic chains between two or three sites at a given distance interval (solid line in Figure 3). Taking the maximum instead the mean number of shared chains (i.e., the mean number of shared tritrophic chains at the given distance interval) captures much better the decrease in similarity with distance (cf. Figure 3 vs. Figure A3).

Previous studies suggest that strong dispersal limitation induces a steep decay in community similarity (Nekola and Peter, 1999; Morlon et al., 2008; McClain et al., 2012). Our results show that in addition to the strong dispersal limitation in each metacommunity (i.e., median value for the intensity of emigration rate, $m_R \sim m_H \sim m_P \sim 0.05$), nonrandom dispersal in heterogeneous landscapes leads to a critical distance threshold above which the maximum site similarity decreases much faster (this critical distance in Figure

3 is roughly 200 km, see the sharp corner of the solid line). This threshold is also characterized by the peak in the variance of the distance-decay in the empirical data (Figure 3, bottom panels show such a peak in the empirical data while Figure A4, panel D shows a peak in simulated metacommunities). Neither models with density independent dispersal, nor models that assume homogeneous landscapes predict such a critical threshold (Figure A4).

We also showed that β richness is strongly influenced by habitat similarity. We contrasted distance-decay in site similarity without (Figure 3A) and with (Figure 3C) accounting for habitat similarity.âIn the first case we compared all tritrophic chains in all habitats within a given site, while in the latter case we compared only tritrophic chains in the same habitat. Altogether there were 20 different habitat types (see section "Plant-aphid-parasitoid data"). In our empirical data, beyond 350 km, the 2-site (Figure 3A) and 3-site (Figure 3B) tritrophic chain similarity values without accounting for habitat similarity were near to zero for most pairwise comparisons. However, after accounting for habitat similarity even sites whose distance is 600 km can be highly similar (Figure 3C). Thus, in addition to nonrandom dispersal dynamics in heterogeneous landscapes, association of tritrophic chains driven by the preference of resources, consumers and parasitoids to specific habitat types is an important factor driving β-tritrophic richness. These results suggest that preferential habitat choice predicts lower turnover rates of tritrophic chains with geographic distance which may improve the fit to the empirical observations (Mouquet and Loreau, 2003; Haegeman and Loreau, 2014). We remark, however, that up to 77% of the empirically observed tritrophic chains are still within the limits of our prediction (dotted lines in 3C).

In our analyses, most of the variation in tritrophic richness remains unexplained (Figure 2 middle panels). Only when dispersal is high, models 3 and 4 predict the observed variance in tritrophic richness correctly (Figure 2, black dots, middle panels C, D), but

in this case the predicted values strongly deviate from the empirical value. For metacommunities with resource-consumer-predator dynamics in each patch an increase in the dispersal rates causes strong and synchronous food web fluctuations (Gouhier et al., 2010) (but see (Koelle and Vandermeer, 2005)). Synchrony in local population fluctuations increases variability in the $\alpha-$richness in the metacommunity (Thibaut and Connolly, 2013; Wang and Loreau, 2014). In our models, increasing correlation in dispersal between different trophic levels can lead to such synchronous fluctuations, thus to better predictions to the observed empirical variability of $\alpha-$tritrophic richness in our multi-trophic metacommunity.

A theory of food webs in spatial landscapes is now entering to a predictive stage (McCann et al., 2005; Dale and Fortin, 2010; Cumming et al., 2010; Gouhier et al., 2010; Gravel et al., 2011; Massol et al., 2011; Poisot et al., 2012; Haegeman and Loreau, 2014). A central challenge in this predictive era is to develop methods to infer the processes driving patterns of food webs across broad geographic regions from empirical observations. This may help us to understand the drivers that shape biodiversity patterns and the biogeography of food webs. This article shows that density-dependent and symmetric dispersal in heterogeneous landscapes may add up to other factors to predict the empirical patterns in local food webs (i.e., $\alpha-$tritrophic richness) and to connect the local patterns to turnover rates across geographically distant food webs (i.e., $\beta-$, and $\gamma-$tritrophic richness).

8 Acknowledgments

We thank Nathaniel Holland, Frederic Briand, Julia Astegiano and three anonymous reviewers for valuable suggestions which greatly improved the manuscript. CJM was supported by the Swiss National Science Foundation project 31003A-144162 and partially supported by a Postdoctoral Fellowship at the National Center for Ecological Analysis and

Synthesis, a Center funded by NSF (Grant #DEB-0553768), the University of California, Santa Barbara, and the State of California. CJM also acknowledges the support by Microsoft Research Ltd., Cambridge, United Kingdom. PS and VK acknowledge the institutional support RVO:60077344. FDL is a postdoctoral research fellow from the Fund for Scientific Research (FWO, Flanders).

9 Tables

Concept	Explanation
Homogeneous landscape	The sum of geographic distances from a focal site to all other sites is normally distributed. In our study we used the empirical mean distance between two sites, 145km, and the empirical variance, 50km to generate the normal distance distribution in Figure 1 (red bars).
Heterogeneous landscape	Distance distribution that deviates significantly from a normal distribution (Figure 1, black bars).
Symmetric dispersal	Dispersal probability between two sites is the same in both directions.
Asymmetric dispersal	Dispersal probability between two sites depends on direction of dispersal.
Density-independent immigration	Colonizing a site is independent of species density.
Negative density-dependent immigration	Colonizing a site is a decreasing function of species density.
Density-dependent mortality	Probability to die is a function of species density.
Density-dependent emigration	Probability to leave a site is a function of species density.

Table 1: Glossary of concepts

Symbol	Explanation	Value
ϕ	Resource (R), consumer (H) or parasitoid (P) metacommunity	
$N_i^{k_\phi}$	Abundance of species k of metacommunity ϕ in site i	SV
S_j^ϕ	Number of species in site j of metacommunity ϕ	SV
C_S^N	Sørensen similarity index for N number of sites	SV
$m_{ij}^{k_\phi}$	Dispersal from site j to site i for species k in metacommunity ϕ	SV
m_ϕ	Intensity of emigration rate of metacommunity ϕ	$\mathcal{U}[7 \times 10^{-1}, 10^{-3}]$
J_i^ϕ	Carrying capacity of site i of metacommunity ϕ	m_1,m_3,m_4: 1000, m_2: $\mathcal{N}[1000, 100]$
ν_ϕ	Immigration rate from the regional species pool of metacommunity ϕ	$\mathcal{U}[10^{-2}, 10^{-4}]$
$M_i^{k_\phi}$	Density-dependent mortality of species k in site i and metacommunity ϕ	SV
μ_ϕ^k	Natural mortality of species k in metacommunity ϕ	1
$\alpha^{k_\phi k'_\varphi}$	Mortality rate of species k in metacommunity ϕ by species k' in metacommunity φ due to predation	1
λ_ϕ	Local birth rate of metacommunity ϕ	1 - ν_ϕ - m_ϕ
d_{ij}	Geographical distance between site i and j	ED
N	Number of sites	ED = 302
C_i^{RH}	Connectance in site i for the resource-consumer food web	ED, [0.24, 0.02]
C_i^{HP}	Connectance in site i for the consumer-parasitoid food web	ED, [0.24, 0.02]

Table 2: Symbols used and parameter values. m_1, m_2, m_3 and m_4 refer to models 1 to 4, respectively. SV and ED refer to state variable and empirical data, respectively.

Tolerance	**Bayes factors**		
	$\log(BF_{m_4/m_j})$	$\log(BF_{m_3/m_j})$	$\log(BF_{m_2/m_j})$
$\epsilon = q_{1\%}$			
	$\log(BF_{m_4/m_1}) = 4.2$	$\log(BF_{m_3/m_1}) = 4.3$	$\log(BF_{m_2/m_1}) = 1$
	$\log(BF_{m_4/m_2}) = 3.2$	$\log(BF_{m_3/m_2}) = 3.3$	
	$\log(BF_{m_4/m_3}) = 0.1$		
$\epsilon = q_{3\%}$			
	$\log(BF_{m_4/m_1}) = 5.3$	$\log(BF_{m_3/m_1}) = 4.4$	$\log(BF_{m_2/m_1}) = 0.9$
	$\log(BF_{m_4/m_2}) = 4.3$	$\log(BF_{m_3/m_2}) = 5.4$	
	$\log(BF_{m_4/m_3}) = 0.04$		
$\epsilon = q_{5\%}$			
	$\log(BF_{m_4/m_1}) = 5.9$	$\log(BF_{m_3/m_1}) = 4.9$	$\log(BF_{m_2/m_1}) = 1.2$
	$\log(BF_{m_4/m_2}) = 4.9$	$\log(BF_{m_3/m_2}) = 5.8$	
	$\log(BF_{m_4/m_3}) = 0.02$		

Table 3: This table compares the four models to predict the α-tritrophic richness. Comparison of the four models are based on Bayes factors (BF_{m_k/m_j}) according to the Jeffreys' scale (decisive: $\log(BF_{m_k/m_j}) > 2$, strong: $1 < \log(BF_{m_k/m_j}) < 2$, substantial: $0.5 < \log(BF_{m_k/m_j}) < 1$, weak: $0 < \log(BF_{m_k/m_j}) < 0.5$). Models in the Jeffreys' scale are denoted by m_k, $k = 2, 3, 4$ and m_j, $j = 1, 2, 3$. The four models are: m_1: Density- and site-independent dispersal dynamics; m_2: Site-dependent dispersal dynamics; m_3: Density-dependent dispersal dynamics; m_4: Density-dependent and asymmetric dispersal dynamics. The three tolerance threshold values, ϵ, correspond to the 1%, 3%, and 5% upper quantile of the distribution of misfits.

Tolerance	**Bayes factors**		
	$\log(BF_{m_4/m_j})$	$\log(BF_{m_3/m_j})$	$\log(BF_{m_2/m_j})$
$\epsilon = q_{1\%}$			
	$\log(BF_{m_4/m_1}) = 2.4$	$\log(BF_{m_3/m_1}) = 2.4$	$\log(BF_{m_2/m_1}) = 0.1$
	$\log(BF_{m_4/m_2}) = 2.5$	$\log(BF_{m_3/m_2}) = 2.3$	
	$\log(BF_{m_4/m_3}) = 0.1$		
$\epsilon = q_{3\%}$			
	$\log(BF_{m_4/m_1}) = 3.5$	$\log(BF_{m_3/m_1}) = 2.8$	$\log(BF_{m_2/m_1}) = 0.6$
	$\log(BF_{m_4/m_2}) = 2.9$	$\log(BF_{m_3/m_2}) = 3.4$	
	$\log(BF_{m_4/m_3}) = 0.16$		
$\epsilon = q_{5\%}$			
	$\log(BF_{m_4/m_1}) = 4.1$	$\log(BF_{m_3/m_1}) = 3.3$	$\log(BF_{m_2/m_1}) = 0.6$
	$\log(BF_{m_4/m_2}) = 3.5$	$\log(BF_{m_3/m_2}) = 3.9$	
	$\log(BF_{m_4/m_3}) = 0.1$		

Table 4: This table compares the four models to predict the γ-tritrophic richness. Meaning of symbols is the same as in Table 3.

Figure Legends

Figure 1. This figure shows distribution of total geographic distances for all sites. The total geographic distance for a given site is the sum of all distances between this focal site and all other sites. The distribution for sampled sites is shown in black, while the normal distribution with the same mean (145km), and variance (50km) is shown in red. This figure shows that the empirical site distribution corresponds to a heterogeneous landscape, as it significantly deviates from the normal site distribution.

Figure 2. This figure shows fit between model and data for α- and γ-tritrophic richness. Each dot represents the misfits (ρ, equation (A-2), y-axis) between a single simulation run of the corresponding model 1-4 and the true value for the mean of α-tritrophic richness ($x-$axis, left) and the γ-tritrophic richness ($x-$axis, right). The variance of α-tritrophic richness predicted by simulations is shown in the middle column. Horizontal line represents the upper quantile 1% of the misfits (with the corresponding tolerance value ϵ_1 = -1338.5) after 100000 replicates for each model. Vertical line in left and right panels represent the mean α-tritrophic richness (16.4, left) and the γ-tritrophic richness (1304, right). Line in the middle panels represent the empirical variance (middle column) of the α-tritrophic richness. Orange, red and black dots represent low ($m_\phi \in [10^{-3}, 10^{-1}]$), medium ($m_\phi \in [10^{-1}, 2 \times 10^{-1}]$), and high ($m_\phi \in [2 \times 10^{-1}, 7 \times 10^{-1}]$) intensity of emigration rate, respectively.

Figure 3. This figure shows empirical distance-decay of maximum similarity (dots, top panels), and empirical distance-decay of variance of similarity (dots, bottom panels) for 2-site β-tritrophic similarity **(A)**, 3-site β-tritrophic similarity **(B)**, and 2-site habitat β-tritrophic similarity **(C)**. We use all tritrophic chains found in each pair (triple) of sites to compute the 2-site (3-site) β-tritrophic similarity in **A (B)**. For the 2-site habitat β-tritrophic similarity in **(C)** we compare only those tritrophic chains within a given habitat type. Solid lines are the mean maximum similarity predicted by model 3 with

symmetric and density-dependent dispersal dynamics together with confidence intervals (dotted lines).

10 Figures

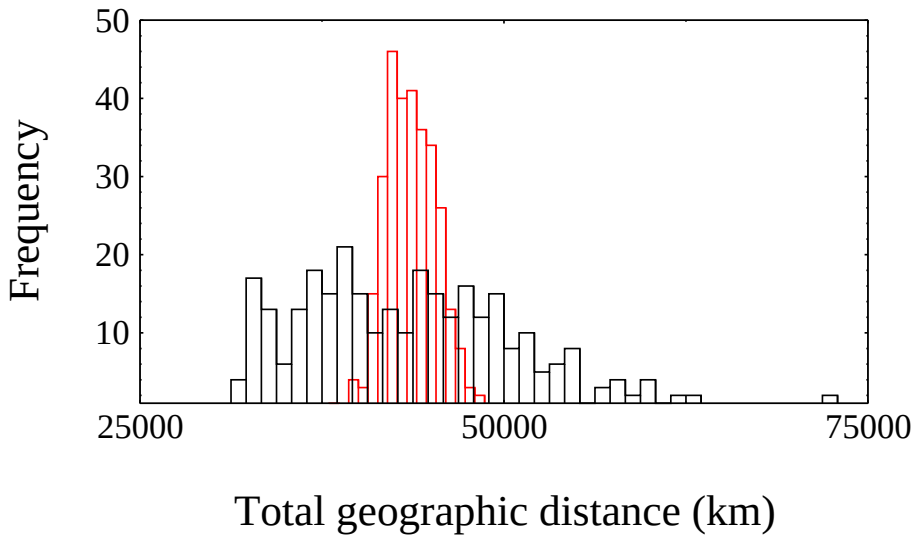
50
40
30
20
10
Frequency
25000
50000
75000
Total geographic distance (km)

Figure 1:

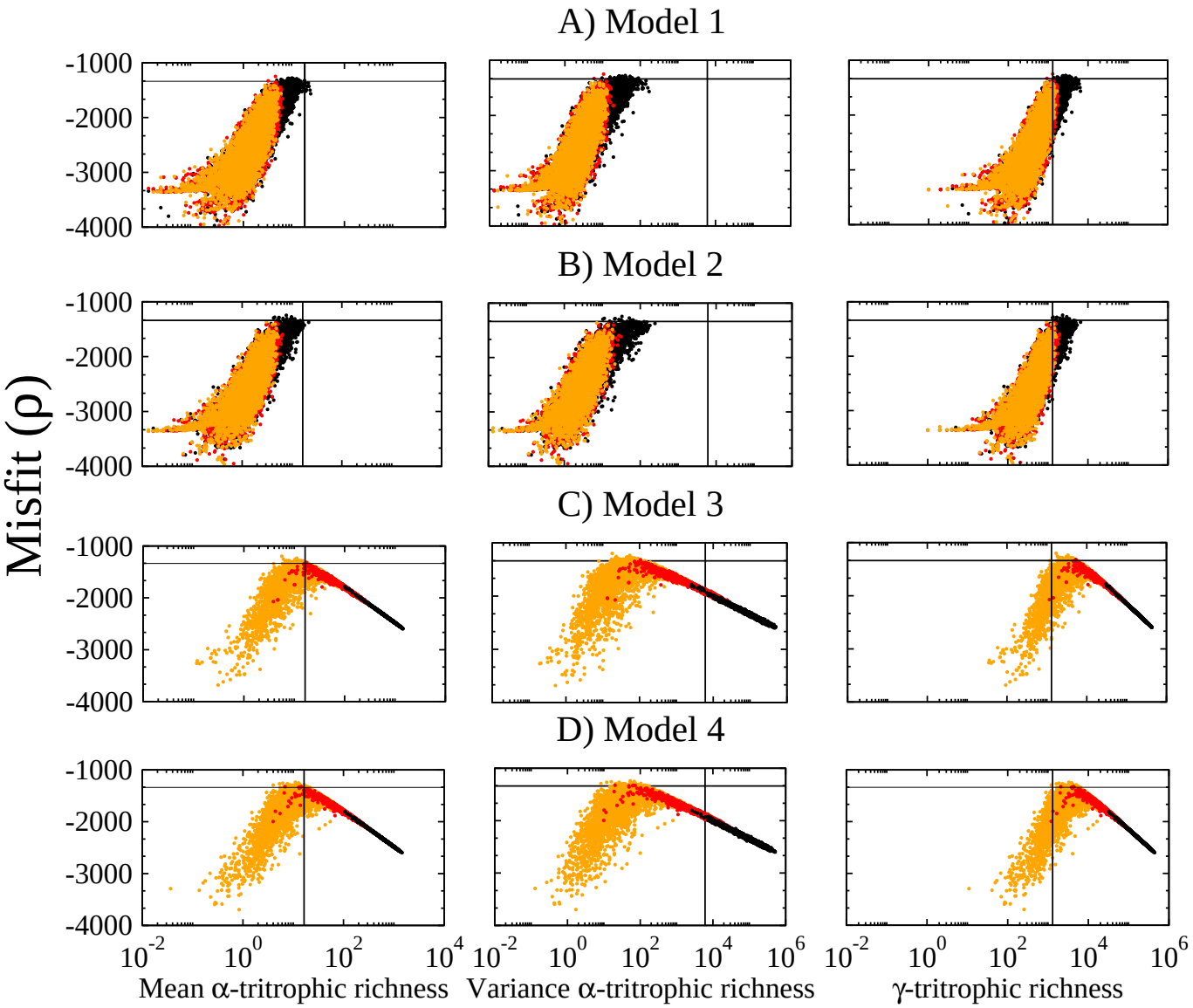
A) Model 1
B) Model 2
C) Model 3
D) Model 4
Misfit (ρ)
-1000
-2000
-3000
-4000
10^-2
10^0
10^2
10^4
10^6
Mean α-tritrophic richness
Variance α-tritrophic richness
γ-tritrophic richness

Figure 2:

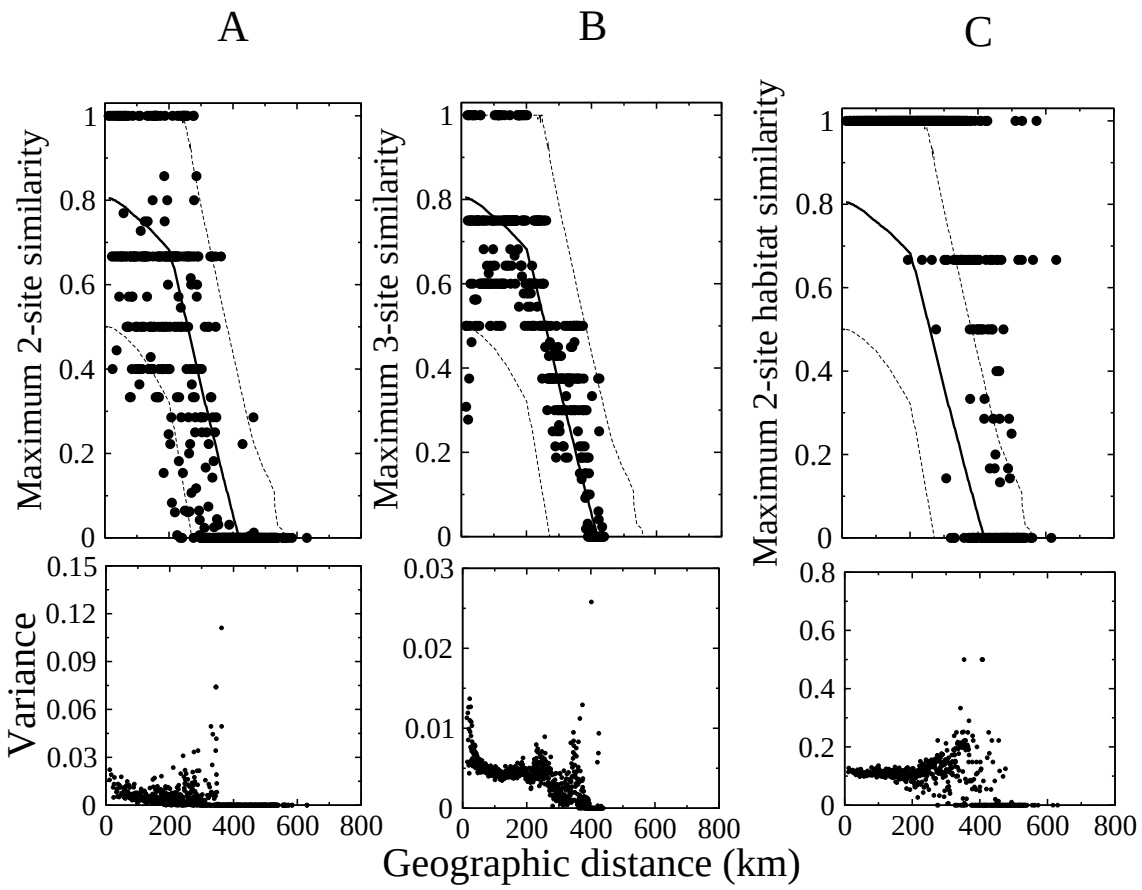
A
B
C
Maximum 2-site similarity
Maximum 3-site similarity
Maximum 2-site habitat similarity
Variance
Geographic distance (km)
0
0.2
0.4
0.6
0.8
1
0.03
0.06
0.09
0.12
0.15
0.01
0.02
200
400
600
800

Figure 3:

Appendix 1 from C. Melián, V. Křivan, F. Altermatt, P. Starý, L. Pellissier, and F. De Laender, Multi-trophic metacommunities and the biogeography of food webs

Multi-trophic metacommunity dynamics

Here, we explain in detail how we combine dispersal with population dynamics in each site. The following equations provide the population dynamics in each site in a multi-trophic metacommunity dynamics context

$$P\left(N_i^{k_\phi}-1|N_i^{k_\phi}\right)=M_i^{k_\phi}\left[\sum_{j=1,j\neq i}^{N}\sum_{k'_\phi=1,k'_\phi\neq k_\phi}^{S_j^\phi} m_{ij}^{k'_\phi}\left(\frac{N_j^{k'_\phi}}{J_j^\phi}\right)+\lambda_\phi\left(\frac{J_i^\phi-N_i^{k_\phi}}{J_i^\phi-1}\right)+\nu_\phi\right]$$

$$P\left(N_i^{k_\phi}+1|N_i^{k_\phi}\right)=(1-M_i^{k_\phi})\left[\sum_{j=1,j\neq i}^{N} m_{ij}^{k_\phi}\left(\frac{N_j^{k_\phi}}{J_j^\phi}\right)+\lambda_\phi\left(\frac{N_i^{k_\phi}}{J_i^\phi-1}\right)+\nu_\phi\right], \tag{A-1}$$

where $M_i^{k_\phi}$ is a compound term that describes density-dependent mortality rate of species k_ϕ in site i and metacommunity ϕ. This mortality is either the natural per capita mortality rate described in this article by $\mu^{k_\phi}\frac{N_i^{k_\phi}}{J_i^\phi}$, or it is the mortality rate due to predation by some consumer or parasitoid species k'_Φ in metacommunity Φ, in which case it is modeled as $\alpha^{k_\phi k'_\Phi}\frac{N_i^{k_\phi}N_i^{'k_\Phi}}{J_i^\phi J_i^\Phi}$. $N_i^{k_\phi}$ and J_i^ϕ are the number of individuals of species k_ϕ in site i of metacommunity ϕ and the number of individuals in site i in metacommunity ϕ, respectively. $\mathcal{S}_j^\phi$ is the total number of species in site j of metacommunity ϕ. In addition to the mortality rate parameters, there are three more metacommunity specific parameters: λ_ϕ, the local birth rate, m_ϕ, the emigration rate, and ν_ϕ, the immigration rate from the regional species pool.

The first equation in (A-1) gives the transition probability for the k_ϕ^{th} species of metacommunity ϕ to decline in abundance by one individual in site i. For this to happen, an individual must die in the k_ϕ^{th} species of metacommunity ϕ, which occurs at a rate

given by $M_i^{k_\phi}$. The first probability inside the brackets is that of an immigration event of some species other than k_ϕ from a site different to i. The second term represents the probability of having a local birth in a species other than k_ϕ with the -1 subtracted in the denominator after the death in the previous step of one individual in this site. The third term describes the probability of an immigration event from the regional species pool. The second equation in (A-1) describes the transition probability for the k_ϕ^{th} species to increase by one individual. For this to happen, there must be no local death in species k_ϕ which is given by $1 - M_i^{k_\phi}$. The other terms in brackets stand for dispersal (the first term), local birth of an individual of species k_ϕ (second term), and immigration of a new species k_ϕ in metacommunity ϕ from the regional species pool. This last event can occur only when there was no such species, i.e., when $N_i^{k_\phi} = 0$ at time $t - 1$.

Appendix 2 from C. Melián, V. Křivan, F. Altermatt, P. Starý, L. Pellissier, and F. De Laender, Multi-trophic metacommunities and the biogeography of food webs

Approximate Bayesian computation: model-data fitting and comparison

To compare the models with the empirical data we use the following algorithm (Grelaud et al., 2009)

1. Generate a replicate from model m^*
2. Generate $\theta^*_{m^*}$ from the prior with parameter values for the emigration rate and the immigration rate from the regional species pool for each plant, aphid and parasitoid metacommunity chosen from a uniform distribution, $\theta^* = [m_R,\ m_H,\ m_P,\ \nu_R,\ \nu_H,\ \nu_P]$,
3. Generate the $\alpha-$, $\beta-$, and $\gamma-$tritrophic richness from the model m^* given the parameter values from θ^*
4. Compute the misfit, $\rho(x_o, x_*)$, between empirical (x_o) and simulated data (x_*) as

$$\rho(x_o^{(1)}, ..., x_o^{(N)}|\nu_\phi, m_\phi) = \sum_{i=1}^{N} \ln(P(x_o^{(i)}|x_*^{(i)}), \nu_\phi, m_\phi), \tag{A-2}$$

where N is the number of sampled sites, $x_o^{(i)}$ and $x_*^{(i)}$ are the empirical and predicted numbers of tritrophic chains in site i, respectively. $P(x_o^{(i)}|x_*^{(i)})$ was calculated as (Tarantola, 2006)

$$P(x_o^{(i)}|x_*^{(i)}) = \frac{1}{x_*^{(i)}(2-e^{-1})} \exp\left(-\left|\frac{x_o^{(i)} - x_*^{(i)}}{x_*^{(i)}}\right|\right) \tag{A-3}$$

The normalization term $\frac{1}{x_*^{(i)}(2-e^{-1})}$ is a correction for the sampling bias, as some sites were sampled less often (e.g., only once) than other sites. In most sites that were

sampled only once only a single tritrophic chain was observed. This is likely due to the sampling effort and the normalization factor corrects for it by penalizing less those sites where several tritrophic chains were observed. We note that when (A-3) is substituted to (A-2), the resulting value ρ will be negative. Results presented in this article are qualitatively the same (Table A1) when the term $\frac{1}{x_*^{(i)}(2-e^{-1})}$ is omitted from formula (A-3).

5. Accept $(\theta^*_{m^*},\ m^*)$ if $\rho(x_o, x_*) > \epsilon$ and the predicted values are within the $\pm$ 0.05 interval of the empirical values of $\alpha-$, and $\gamma-$tritrophic richness, otherwise, start again in 1.

Once a sample of S values was obtained for each model, the standard Monte Carlo approximation of the posterior probabilities based on the number of times each model was chosen, namely $\mathbb{P}(\mathcal{M} = \mathrm{m}|x_o) = \#(m^{i*} = m)/S$, where $\#(m^{i*} = m)$ denotes the number of simulated m^{i*}'s equal to m. The Bayes factor associated with the evidence provided by the empirical data x_o in favor of model k relative to model j is as

$$BF_{m_k/m_j} = \frac{1 + \#(m^{i*} = m_k)}{1 + \#(m^{i*} = m_j)} \times \frac{\pi(\mathcal{M} = m_j)}{\pi(\mathcal{M} = m_k)}, \tag{A-4}$$

and the empirical frequencies of visits to the model $\mathbb{P}(\mathcal{M} = m|x_o)$ will depend on the tolerance threshold, ϵ, corresponding in our study to the 1%, 3%, and 5% quantile of the misfits (Table 3 and 4). To compare model predictions with the empirical values of the β-tritrophic richness the mean and confidence interval (CI) were generated by taking the percentiles $5th$ and $95th$ from 3% quantile values (Figure 3).

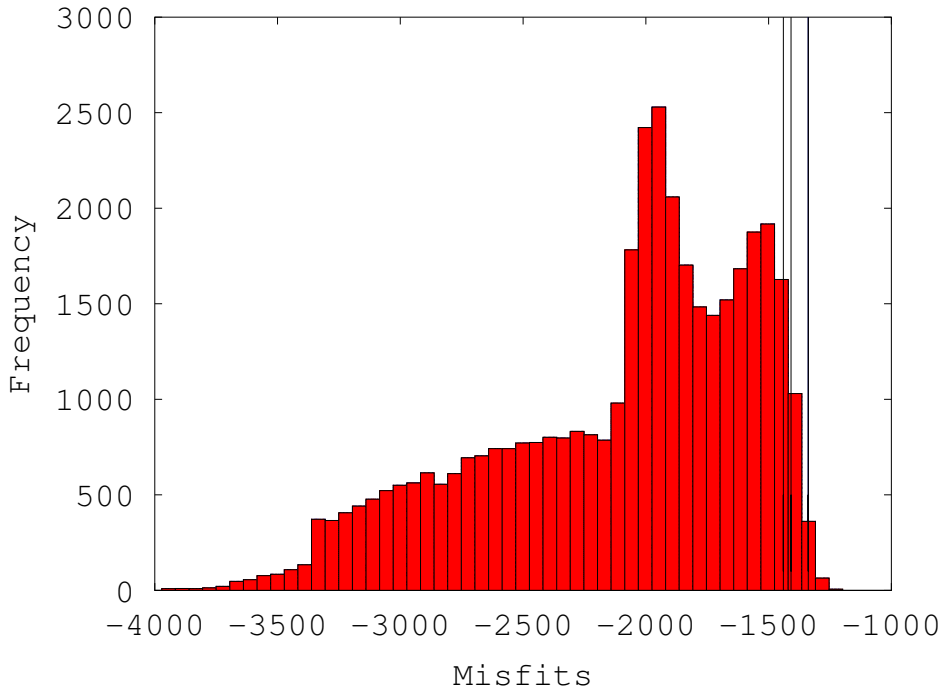

Figure A1: This figure shows the distribution of misfits by comparing the empirical data and the simulated data. Vertical red lines show the three tolerance threshold values, ϵ, corresponding to the 1%, 3%, and 5% upper quantile of the misfits.

Tolerance	**Bayes factors**		
	$\log(BF_{m_4/m_j})$	$\log(BF_{m_3/m_j})$	$\log(BF_{m_2/m_j})$
$\epsilon = q_{1\%}$			
	$\log(BF_{m_4/m_1}) = 2.3$	$\log(BF_{m_3/m_1}) = 2.3$	$\log(BF_{m_2/m_1}) = 0.4$
	$\log(BF_{m_4/m_2}) = 2.1$	$\log(BF_{m_3/m_2}) = 2.1$	
	$\log(BF_{m_4/m_3}) = 0.3$		
$\epsilon = q_{3\%}$			
	$\log(BF_{m_4/m_1}) = 2.4$	$\log(BF_{m_3/m_1}) = 2.4$	$\log(BF_{m_2/m_1}) = 0.3$
	$\log(BF_{m_4/m_2}) = 1.9$	$\log(BF_{m_3/m_2}) = 3.4$	
	$\log(BF_{m_4/m_3}) = 0.4$		
$\epsilon = q_{5\%}$			
	$\log(BF_{m_4/m_1}) = 2.9$	$\log(BF_{m_3/m_1}) = 2.9$	$\log(BF_{m_2/m_1}) = 1.2$
	$\log(BF_{m_4/m_2}) = 2.1$	$\log(BF_{m_3/m_2}) = 2.5$	
	$\log(BF_{m_4/m_3}) = 0.5$		

Table A1: This table compares the four models to predict the $\alpha-$tritrophic richness using only the absolute misfits. Comparison of the four models are based on Bayes factors (BF_{m_k/m_j}) according to the Jeffreys' scale (decisive: $\log(BF_{m_k/m_j}) > 2$, strong: $1 < \log(BF_{m_k/m_j}) < 2$, substantial: $0.5 < \log(BF_{m_k/m_j}) < 1$, weak: $0 < \log(BF_{m_k/m_j}) < 0.5$). Models in the Jeffreys' scale are denoted by m_k, $k = 2, 3, 4$ and m_j, $j = 1, 2, 3$. The four models are: m_1: Density- and site-independent dispersal dynamics; m_2: Site-dependent dispersal dynamics; m_3: Density-dependent dispersal dynamics; m_4: Density-dependent and asymmetric dispersal dynamics. The three tolerance threshold values, ϵ, correspond in this case to the 1%, 3%, and 5% lower quantile of the distribution of misfits.

Appendix 3 from C. Melián, V. Křivan, F. Altermatt, P. Starý, L. Pellissier, and F. De Laender, Multi-trophic metacommunities and the biogeography of food webs

Sampling effort

We tested the robustness of $\gamma-$ (Figure A2, panels A, B), $\alpha-$ (Figure A2, panels C, D) and $\beta-$ (Figure A2, panels E, F) tritrophic richness to sampling effort by studying how these indexes change when the number of sample sites increases (Polis, 1991; Bersier and Sugihara, 1999). This was done by randomly sampling an increasing number of sites taking into account all the tritrophic chains observed in each site, starting by 5, 10, 20, 50, 100, 200, and 300 sites from the original dataset and calculating all three richness measures for each of these subsets. Because these measures were quantified as distributions, we described each measure using its mean value and its standard deviation. Random samplings were repeated 1000 times to account for variability in tritrophic richness among sites.

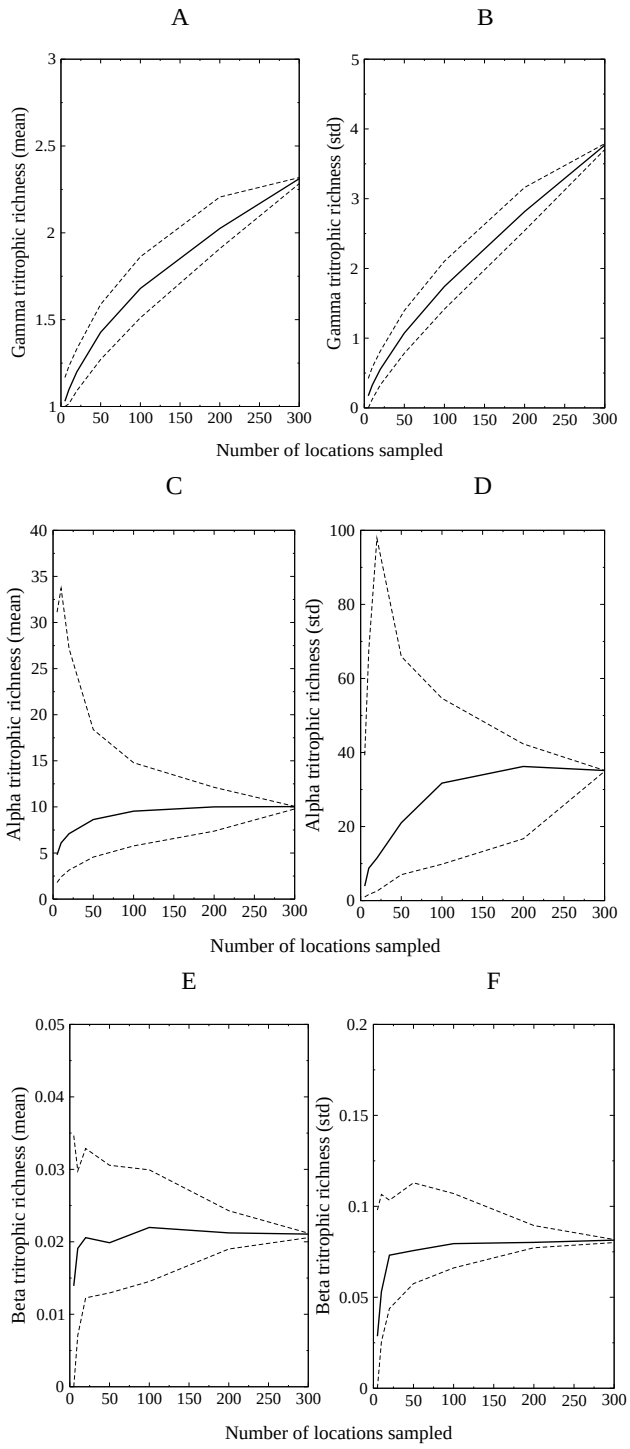

Figure A2: This figure shows the effect of sampling effort on the mean and standard deviation of the $\gamma-$ (A-B), $\alpha-$ (C-D), and $\beta-$tritrophic richness (E-F). Solid and dashed lines represent the mean, $5th$ and $95th$ percentiles after 1000 iterations.

Appendix 4 from C. Melián, V. Křivan, F. Altermatt, P. Starý, L. Pellissier, and F. De Laender, Multi-trophic metacommunities and the biogeography of food webs

Distance-decay

We calculated the distance-decay of mean similarity in heterogeneous landscapes for model 1 (red, Figure A3) and model 3 (black, Figure A3). The distance-decay pattern when using the mean similarity is not as clear as the pattern when using the maximum similarity (cf. Figure 3 vs. Figure A3). Model predictions of the distance-decay using the maximum similarity and the variance are shown in Figure A4. These results show that predictions for the model 1 (Figure A4, panels A and C, red homogeneous and black heterogeneous distance distribution) are different from predictions of model 3 (Figure A4, panels B and D, red homogeneous and black heterogeneous distance distribution). The predicted distance-decay of variance of similarity shows no peak across all the geographic distances for the model 1 with density-independent dispersal neither in homogeneous (Figure A3, panel C, red dots) nor in heterogeneous (panel C, black dots) landscapes. Model 3 predicts a threshold when applied to a heterogeneous landscape defined as the distance at which a peak in the variance of the distance-decay occurs (Figure A4, panel D). We remark that predictions for model 2 are similar to model 1 and predictions for model 4 are similar to model 3.

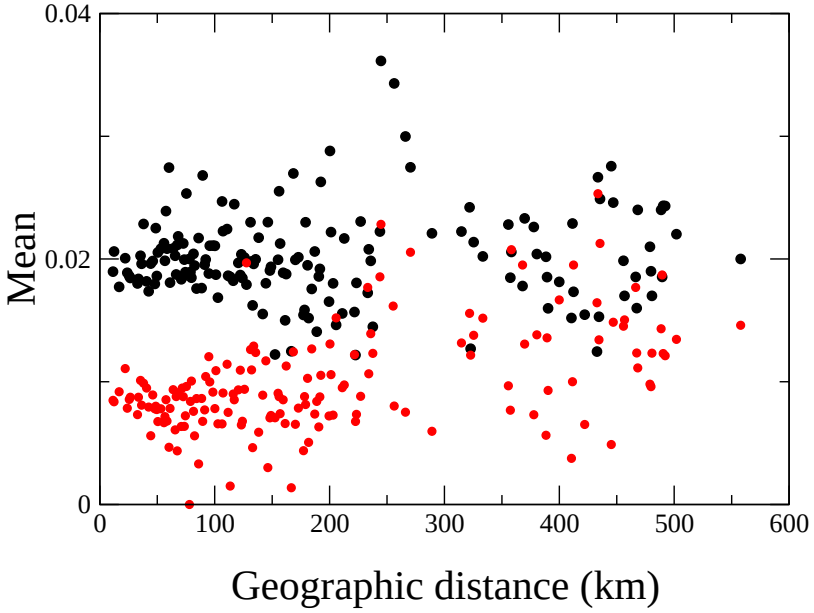

Figure A3: This figure shows empirical distance-decay of mean similarity (model 1, red, and model 3, black). For each replicate we calculate the mean for all the pairwise comparisons at a given geographic distance and each dot represents the mean after 10000 replicates using migration values, m, from a uniform distribution with range [0.001,0.1].

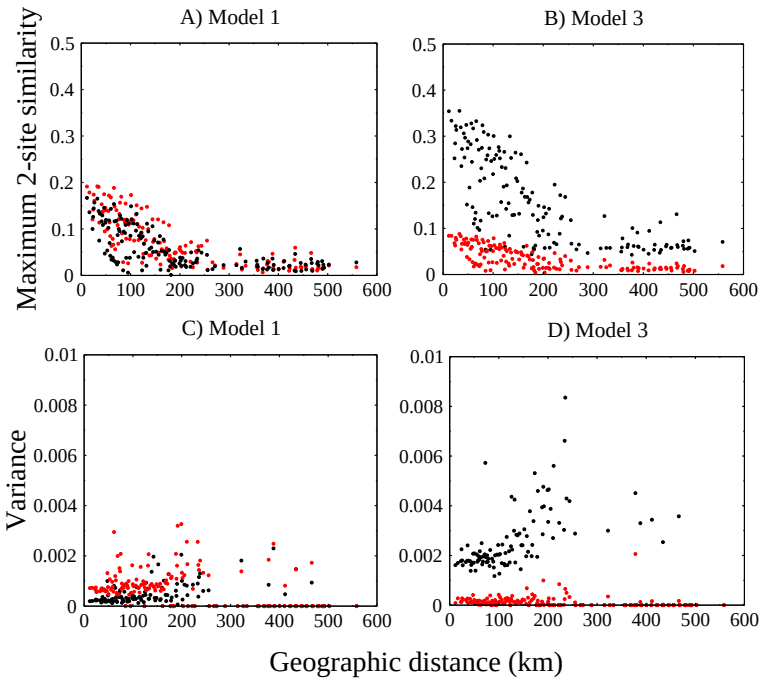

Figure A4: **A)** Distance-decay for the maximum 2-site similarity for the model with symmetric and density-independent dispersal dynamics. **B)** Distance-decay for the maximum 2-site similarity for the model with symmetric and density-dependent dispersal dynamics. **C)** Variance-decay of similarity with distance for the 2-site similarity for the model with symmetric and density-independent dispersal dynamics. **D)** Variance-decay of similarity with distance for the 2-site similarity for the model with symmetric and density-dependent dispersal dynamics. Red dots (black dots) assume homogeneous (heterogeneous) landscapes. Each dot represents mean values after 10000 replicates using dispersal values, m, in the range $[0.001, 0.1]$.

References

1. Amarasekare, P. 2008. Spatial dynamics of foodwebs. Annual Review Ecology, Evolution and Systematics 39:479–500.

2. Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. Global Ecology and Biogeography 16:743–753.

3. Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. Annual Review Ecology, Evolution and Systematics 41:379–406.

4. Bersier, L.-F., D. P. and Sugihara, G. 1999. Scale-invariant or scale-dependent behavior of the link density property in food webs. The American Naturalist 153:676–682.

5. Bolker, B., 2004. Continuous-space models for population dynamics. In: Ecology, Genetics, and Evolution of Metapopulations (eds Hanski, I., & Gaggioti, O. E.). Elsevier Science, San Diego, pp. 45-69.

6. Boulangeat, I., Gravel, D., and Thuiller, W. 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. Ecology Letters 15:584–593.

7. Briggs, C. J. and Hoopes, M. F. 2004. Stabilizing effects in spatial parasitoid-host and predator-prey models: a review. Theoretical Population Biology 65:299–315.

8. Carrara, F., Altermatt, F., Rodriguez-Iturbe, I., and Rinaldo, A. 2012. Dendritic connectivity controls biodiversity patterns in experimental metacommunities. Proceedings of the National Academy of Sciences of the USA 109:5761–5766.

9. Channell, R. and Lomolino, M. V. 2000. Dynamic biogeography and conservation of endangered species. Nature 403:84–86.

10. Cumming, G. S., Bodin, O., Ernstson, H., and Elmqvist, T. 2010. Network analysis in conservation biogeography: challenges and opportunities. Diversity and Distributions 16:414–425.

11. Dale, M. and Fortin, M.-J. 2010. From graphs to spatial graphs. Annual Review Ecology, Evolution and Systematics 41:21–38.

12. Diserud, O. H. and Ødegaard, F. 2007. A multiple-site similarity measure. Biology letters 3:20–22.

13. Dunne, J. A., 2006. The network structure of food webs. In: Ecological Networks: Linking Structure to Dynamics in Food Webs (eds. Pascual, M. Dunne, J. A.).

14. Economo, E. P. and Keitt, T. 2008. Species diversity in neutral metacommunities: a network approach. Ecology Letters 11:52–62.

15. Fretwell, D. S. and Lucas, H. L. 1970. On territorial behavior and other factors influencing habitat distribution in birds. Acta Bio-theoretica 19:16–32.

16. Gouhier, T. C., Guichard, F., and Gonzalez, A. 2010. Synchrony and stability of food webs in metacommunities. The American Naturalist 175:E16–E34.

17. Gravel, D., Massol, F., Canard, E., Mouillot, D., and Mouquet, N. 2011. Trophic theory of island biogeography. Ecology Letters 14:1010–1016.

18. Grelaud, A., Robert, C. P., Marin, J.-M., Rodolphe, F., and Taly, J.-F. 2009. ABC likelihood-free methods for model choice in Gibbs random fields. Bayesian Analysis 4:317–336.

19. Haegeman, B. and Loreau, M. 2014. General relationships between consumer dispersal, resource dispersal and metacommunity diversity. Ecology Letters 17:175–184.

20. Hanski, I. 1999. Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. Oikos 87:209–219.

21. Holt, R. D. and Hoopes, M., 2005. Food Web Dynamics in a Metacommunity Context: Modules and Beyond. In: Metacommunities: Spatial Dynamics and Ecological Communities (eds Holyoak, M., Leibold, M. A., & Holt, R. D.). The University of Chicago Press, Chicago, pp. 68-93.

22. Huffaker, C. B. 1958. Experimental studies on predation: dispersion factors and predator-prey oscillations. Hilgardia: A Journal of Agricultural Science 27:795–834.

23. Kissling, W. D., Dormann, C. F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G. J., Montoya, J. M., Römermann, C., Schiffers, K., Schurr, F. M., Singer, A., Svenning, J.-C., Zimmermann, N. E., and O'Hara, R. B. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. Journal of Biogeography 39:2163–2178.

24. Koelle, K. and Vandermeer, J. 2005. Dispersal-induced desynchronization: from metapopulations to metacommunities. Ecology Letters 8:167–175.

25. Křivan, V., 2008. Predator-prey models. Pages 2929-2940 in S. E. Jørgensen and B. D. Fath, eds. Population Dynamics. Vol. 4 of Encyclopedia of Ecology, 5 vols. Elsevier, Oxford, UK.

26. Křivan, V. 2014. Competition in di- and tri-trophic food web modules. Journal of Theoretical Biology 343:127–137.

27. Křivan, V., Cressman, R., and Schneider, C. 2008. The ideal free distribution: A review and synthesis of the game theoretic perspective. Theoretical Population Biology 73:403–425.

28. Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., and Hoopes, M. F. 2004. The metacommunity concept: a framework for multi-scale community ecology. Ecology Letters 7:601–613.

29. Levins, R. 1962. Theory of fitnesses in a heterogeneous environment, I. the fitness set and adaptive function. The American Naturalist 96:361–378.

30. Magurran, A., 2004. Measuring biological diversity. Blackwell Publishing, Oxford, UK.

31. Massol, F., Gravel, D., Mouquet, N., Cadotte, M. W., Fukami, T., and Leibold, M. A. 2011. Linking community and ecosystem dynamics through spatial ecology. Ecology Letters 14:313–323.

32. McCann, K. S., Rasmussen, J. B., and Umbanhowar, J. 2005. The dynamics of spatially coupled food webs. Ecology Letters 8:513–523.

33. McClain, C. R., Stegen, J. C., and Hurlbert, A. H. 2012. Dispersal, enrironmental niches and oceanic-scale turnover in deep-sea bivalves. Proceedings of the Royal Society B 279:1993–2002.

34. Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R., and Green, J. L. 2008. A general framework for the distance-decay of similarity in ecological communities. Ecology Letters 11:904–917.

35. Mouquet, N. and Loreau, M. 2003. Community patterns in source-sink metacommunities. The American Naturalist 162:544–557.

36. Murdoch, W. W., Briggs, C. J., and Nisbet, R. M., 2003. Consumer-Resource Dynamics. Princeton Univ. Press, Princeton, USA.

37. Nekola, J. C. and Peter, S. 1999. The distance decay of similarity in biogeography and ecology. Journal of Biogeography 26:867–878.

38. Pillai, P., Gonzalez, A., and Loreau, M. 2011. Metacommunity theory explains the emergence of food web complexity. Proceedings of the National Academy of Sciences of the United States of America 108:19293–19298.

39. Poisot, T., Canard, E., Mouillot, D., Mouquet, N., and Gravel, D. 2012. The dissimilarity of species interaction networks. Ecology Letters 15:1353–1361.

40. Polis, G. A. 1991. Complex trophic interactions in deserts: An empirical critique of food web theory. The American Naturalist 138:123–155.

41. Rezende, E. L., Albert, E. M., Fortuna, M. A., and Bascompte, J. 2009. Compartments in a marine food web associated with phylogeny, body mass, and habitat structure. Ecology Letters 12:779–788.

42. Shurin, J. 2001. Interactive effects of predation and dispersal on zooplankton communities. Ecology 82:3404–3416.

43. Smith, D., Bailey, J. K., Shuster, S. M., and Whitham, T. G. 2011. A geographic mosaic of trophic interactions and selection: trees, aphids and birds. Journal of Evolutionary Biology 24:422–429.

44. Starý, P., 2006. Aphid parasitoids of the Czech Republic (Hymenoptera: Braconidae, Aphidiinae). ACADEMIA, Praha.

45. Tarantola, A. 2006. Popper, Bayes and the inverse problem. Nature physics 2:492–494.

46. Thibaut, L. and Connolly, S. 2013. Understanding diversityâstability relationships: towards a unified model of portfolio effects. Ecology Letters 16:140–150.

47. Thuiller, W., Münkemüller, T., Lavergne, S., Mouillot, D., Mouquet, N., Schiffers, K., and Gravel, D. 2013. A road map for integrating eco-evolutionary processes into biodiversity models. Ecology Letters 16:94–105.

48. Vanpeteghem, D. and Haegeman, B. 2010. An analytical approach to spatio-temporal dynamics of neutral community models. Journal Mathematical Biology 61:323–357.

49. Wang, S. and Loreau, M. 2014. Ecosystem stability in space: α, β and γ variability. Ecology Letters 8:1175–1182.